ABSTRACT

The present invention provides an approach for identifying with high accuracy, a known protein or a variant of the known protein derived from the same genomic gene as in a target protein to be analyzed, based on a mass spectrometric result of a plurality of peptide fragments obtained from site-specific enzymatic digestion of the target protein to be analyzed by referring to nucleotide sequences of genes encoding known proteins on a database and to deduced full-length amino acid sequences thereof. In the approach of the present invention, a candidate known protein of identification is specified with high accuracy by such a process comprising steps of comparing actually measured molecular weight values of the peptide fragments derived from the target protein to be analyzed, which are obtained with the use of peptide fragmentation by the site-specific proteolytic treatment, with predicted molecular weight values of peptide fragments predicted from the deduced full-length amino acid sequences of the known proteins and making comparison in terms of the numbers of matching fragments, the consecutiveness of amino acid sequences of matching fragments of the known protein, and the prediction of variation in a mismatching fragment.